GENE CLUSTER

BACKGROUND OF THE INVENTION

Polyketides constitute a large and highly diverse group of secondary metabolites synthesized by bacteria, fungi and plants, with a broad range of biological activities and medical applications. They include anti-cancer agents (Daunorubicin), antibiotics (tetracyclines, erythromycin etc.), immunosuppressants (macrolide FK506) and compounds with mycotoxic activity (aflatoxins, ochratoxins, ergochromes, patulin etc.). Polyketides are synthesized by repetitive condensations of acetate or propionate monomers in a similar way to that of fatty acid biosynthesis. Structural diversity of polyketides is achieved through different thioester primers, varying chain extension units used by the polyketide synthases (PKSs), and variations in the stereochemistry and the degree of reduction of intermediates. Diversity is also achieved by subsequent processing, such as alkylations, oxidations, O-methylations, glycosylations and cyclizations. Genetic studies indicated that gene organization of functional units and motif patterns of various PKSs are similar. This similarity was used to identify and obtain new PKS systems in both gram negative and gram positive bacteria.

PKS systems are classified into two types: type I PKSs are large, multifunctional enzymes, containing a separate site for each condensation or modification step. These represent "modular PKSs" in which the functional domains

encoded by the DNA sequence are usually ordered parallel to the sequence of reactions carried out on the growing polyketide chain. Type II PKSs are systems made up of individual enzymes, in which each catalytic site is used repeatedly during the biosynthetic process.

Genetic studies on prokaryotic PKSs have focused on gram positive microorganisms, particularly on actinomycetes. Myxobacteria are gram negative bacteria that produce a large number of secondary metabolites, including polyketides. *Myxococcus xanthus* produces TA (Rosenberg, et al., 1973; Rosenberg, et al., 1984), which is an antibacterial antibiotic.

The polyketide antibiotic Tel-Aviv (hereinafter TA) (Rosenberg, et al., 1973) is synthesized by the gram negative bacterium *Myxococcus xanthus* in a unique multi-step process incorporating a glycine molecule into the polyketide carbon chain, which is elongated through the condensation of 11 acetate molecules by a type I polyketide synthase (PKSs).

The antibiotic TA was crystallized and its chemical properties were determined. It is a macrocyclic polyketide synthesized through the incorporation of acetate, methionine, and glycine. It inhibits cell wall synthesis by interfering with the polymerization of the lipid-disaccharide-pentapeptide and its ability to adhere avidly to tissues and inorganic surfaces makes it potentially useful in a wide range of clinical applications, such as treating gingivitis.

A growing interest in the study of PKS systems and peptide synthetase systems stems from the need to develop new potent biologically active compounds. The use of combinatorial genetics in both systems (PKS and peptide synthetase) separately has led to the production of new polyketides and new peptides.

It would therefore by useful to be able to generate new biological agents from secondary metabolites of the antibiotic TA.

SUMMARY OF THE INVENTION

According to the present invention, there is provided a purified, isolated and cloned DNA sequence partially encoding a functional portion of a polypeptide component required for the synthesis of antibiotic TA. Also provided are purified, isolated and cloned DNA sequences encoding a polypeptide component required for postmodification of antibiotic TA and encoding a gene product involved in the regulation of the biosynthesis of antibiotic TA. A purified, isolated and cloned DNA sequence having a DNA sequence (Seq. ID No~~s~~ 1 ~~and~~ 2) encoding a polypeptide component required for encoding the TA gene cluster and any mutations thereof is provided. Also provided are methods of using the TA genes for combinatorial genetics and of using the TA genes encoding for synthesis and modification or regulation of antibiotic TA.

DESCRIPTION OF THE DRAWING

Other advantages of the present invention will be readily appreciated as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawing wherein:

Figure 1 shows the physical maps of the DNA regions involved in TA synthesis.

DETAILED DESCRIPTION OF THE INVENTION

The present invention consists of a DNA and Amino acid sequence of at least 42 kb encoding genes involved in TA production and *Myxococcus xanthus* as best shown in Seq. ID No:1 through 17 and cosmid clones containing the entire TA gene DNA sequences. The TA gene cluster has been purified, isolated, and cloned. The purification, isolation and cloning was done according to the methods described in Marshak et al, "Strategies for Protein Purification and Characterization. A laboratory course manual." CSHL Press, 1996.

A DNA fragment of at least 42 kb (Figure 1), encoding genes involved in TA production in *Myxococcus xanthus* has been identified, cloned and analyzed. These steps were done in accordance with Marshak et al, "Strategies for Protein Purification and Characterization. A laboratory course manual." CSHL Press, 1996. This

fragment contains a large region of about 20 kb, encoding the genes responsible for the regulation and the post-modification of TA. An additional fragment of approximately 8-10 kb located 10-20 kb downstream of the post-modification region, encodes the enzyme responsible for the incorporation of the glycine into the polyketide chain. This novel polypeptide is made up of a peptide synthetase unit lying between two PKS modules.

The potential of this unique polypeptide in combining the two systems can lead to a new family of compounds, emerging from various combinations which can be utilized for combinatorial genetics. Such utilization can produce, for example, new bioactive agents, new polyketides and new peptides. Additionally, the TA gene cluster can be utilized in a method for the synthesis, modification or regulation of the TA antibiotic.

Mutations imparting defects into the TA gene cluster can be point mutations, deletions or insertions. The mutations can occur within the nucleotide sequence of the allele of the TA gene cluster such that the resulting amino acid sequence of the TA gene cluster product is altered.

In one embodiment of the present invention, the TA gene cluster can be included in a vector or recombinant expression vector. This vector containing the TA gene cluster is able to transform a suitable eucaryotic or procaryotic host cell. A suitable host cell can be determined by one skilled in the art. An example of a

suitable cell which can be transformed by the TA gene cluster is an E. coli cell.

In another embodiment of the present invention, the a DNA fragment encoding the TA gene cluster can be cloned into a cosmid, as shown in Figure 1. This DNA fragment contains a large region of about 20kb, encoding the genes responsible for the regulation and the post-modification of TA. An additional fragment of approximately eight to ten kb is located 10-20 kb downstream of the post-modification region and encodes the enzyme responsible for the incorporation of the glycine into the polyketide chain. The novel polyketide chain is made up of a peptide synthetase unit lying between two PKS modules (See Figure 1).

The above discussion provides a factual basis for the use of the TA gene cluster. The methods used with and the utility of the present invention can be shown by the following non-limiting examples and accompanying figure.

EXAMPLES

GENERAL METHODS:

METHODS:

General methods in molecular biology: Standard molecular biology techniques known in the art and not specifically described are generally followed as in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Springs Harbor Laboratory, New York (1989, 1992), and in Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Maryland (1989). Polymerase

chain reaction (PCR) is carried out generally as in *PCR Protocols: A Guide To Methods And Applications*, Academic Press, San Diego, CA (1990). Reactions and manipulations involving other nucleic acid techniques, unless stated otherwise, are performed as generally described in Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, and methodology as set forth in United States patents 4,666,828; 4,683,202; 4,801,531; 5,192,659 and 5,272,057 and incorporated herein by reference. In-situ (In-cell) PCR in combination with Flow Cytometry can be used for detection of cells containing specific DNA and mRNA sequences (Testoni et al, 1996, Blood 87:3822.)

Recombinant Protein Purification

Marshak et al, "Strategies for Protein Purification and Characterization. A laboratory course manual." CSHL Press, 1996.

Example 1:

Analysis of the TA gene cluster by chromosomal restriction map.

Chromosomal DNA of several transposition mutants (ER-2514, ER-1037, ER-1030, ER-1311, ER-7513, ER-3708, ER-4639 and ER-6199; Varon *et al.*, 1992) was extracted, digested with restriction enzymes that cut within the transposon, and analyzed by Southern hybridization with six different probes (originating from Tn*V* and Tn5*lac*). We used probes designed to hybridize either to the entire transposon, or to its 5' or 3' ends. A chromosomal restriction map of the whole gene cluster was constructed on the basis of these results (Figure 1). The data refined the transduction

map (Varon *et al.*, 1992) and further indicated that all the genes in the cluster are transcribed in the same direction (see Figure 1).

Preparation of TA-specific probes

DNA from the Tn*V* mutant ER-4639, ER1311 and ER-6199 was digested with *Kpn*I (does not restrict Tn*V*), self-ligated and transformed into *E. coli* XL1-Blue MR using the transposon-derived kanamycin resistance for selection. Tranformant clones pPYT4639, pPYT1311/p5 and pPYT6199 carried a 1.5 kb, 2.3 kb and a 11.2 kb fragment, respectively (see Figure 1).

Cloning of a *M. xanthus* DNA region encoding genes involved in TA biosynthesis.

A library of *M. xanthus* ER-15 was constructed in the cosmid vector SUPERCOS-1 and screened using specific TA probes obtained from transposition mutants (ER-4639, ER-1311 and ER-6199, see map) that contain a Tn*V* transposon. Seventy four recombinant cosmids that carried genes required for TA production were identified through colony hybridization. The cosmids, pPYCC64 and pPYCC44, which hybridized to these probes were further characterized through restriction analysis (see Figure 1) and sub cloned for sequencing.

Throughout this application, various publications, including United States patents, are referenced by author and year and patents by number. Full citations for the publications are listed below. The disclosures of these publications and patents in

their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

The invention has been described in an illustrative manner, and it is to be understood that the terminology which has been used is intended to be in the nature of words of description rather than of limitation.

Obviously, many modifications and variations of the present invention are possible in light of the above teachings. It is, therefore, to be understood that within the scope of the appended claims, the invention may be practiced otherwise than as specifically described.

REFERENCES

1. Rosenberg, E., Vaks, B. and Zuckerberg. A. Bactericidal action of an antibiotic produced by Myxococcus xanthus. Antimicrob. Agents. Chemother. 4:507-513 (1973).

2. Rosenberg, E., Porter, J.M., Nathan, P.N., Manor, A. and Varon, M. Antibiotic TA: an adherent antibiotic. Bio/Technology. 2:796-799 (1984).

3. Varon *et al.*, 1992

4. Marshak et al, "Strategies for Protein Purification and Characterization. A laboratory course manual." CSHL Press, 1996.

5. Testoni et al, 1996, Blood 87:3822.

6. *PCR Protocols: A Guide To Methods And Applications*, Academic Press, San Diego, CA (1990).

7. Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Springs Harbor Laboratory, New York (1989, 1992).

8. Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Maryland (1989).

SEQ LISTING PAGE(s)

REGION 1:

Ta1 - Peptidesynthetase unit-PKS module.

FRAGMENT size(aa):2392

VDPARLTRAWEGLLERYPLLAGAIRVEGTEPVIVPSGQVSAEVHEVPSVSDSALVATLRASAKVPFDLAC
GPLARLHLYSRSEHEHVLLLCFHHLVLDGASVAPLLDALRERYAGTEAKAGLLEVPIVAPYRAAVEWEQ
LAIGGDEGRRHLDYWRHVLATPVPPPLNLPTDRPRSATGLDSEGATHSQRVPTEQALRLREFARAQQVS
LPTVLLGLYYALLHRHTRQDDVVVGIPTMGRPRAELATAIGYFVNVMAVRARGLGQHSFGSLLRHLHDS
VIDGLEHAHYPFPRVVKDLRLSNGPEEAPGFQTMFTFQSLQLTSAPPRPEPRSGGLPELEPLDCVHQEGAY
PLELEVVEGAKGLTLHFKYDARLYEADTVERMARQLLRAADQVADGVESPLSALSWLDDEERRTLLRD
WNATATPFLEDLGVHELFQRQARETPDAMAVSYEGHSLSYQALDTRSREIAAHLKSFGVKPGALVGIYL
DRSAELVAAMLGVLSAGAAYVPLDPVHPEDRLRYMLEDSGVVVVLARQASRDKVAAIAGASCKVCVLE
DVKAGATSAPAGTSPNGLAYVIYTSGSTGRPKGVMIPHRGVVNFLLCMRRTLGLKRTDSLLAVTTYCFD
IAALELLLPLCAGAQVILASAETVRDAQALKRALRTHRPTLMQATPATWTLLFQSGWENAERVRILCGGE
ALPESLKAHFVRTASDVWNMFGPTETTIWSTMAKVSASRPVTIGKPIDNTQVYVLDDRMQPVPIGVPGE
LWIAGAGVACGYLNRPALTAERFVSNPFTPGTTLYRTGDLARWRADGEVEYLGRLDHQVKVRGFRIEM
GEIEAQLAGHPSVKNCAVVAKELNGTSQLVAYCQPAGTSFDEEAIRAHLRKFLPDYMVPAHVFAVDAIP
LSGNGKVDRGQLMARPVVTRRKTSAVHARSPVEATLVELWKNVLQVNEVGVEDRFFEVGGDSVLAAV
LVEEMNRRFDTRLAVTDLFKYVNIRDMARHMEGATAQARTGATEPAREDTASERDYEGSLAVIGISCQL
PGAADPWRFWKNLREGRDSVVAYRHEELRELGVPEEVLRDSRYVAVRSSIEDKECFDPHFFGLTARDAS
FMDPQFRLLLMHAWKAVEDAATTPERLGPCGVFMTASNSFYHQGSPQFPADGQPVLRTAEEYVLWVLA
QAGSIPTMVSYKLGLKGPSLFVHTNCSSSLSALYVAQQAIAAGDCQTALVGAATVFPSANLGYLHQRGL
NFSSAGRVKAFDAAADGMIAGEGVAVLVVKDAAAAVRDGDPIYCLVRKVGINNDGQDKVGLYAPSAT
GQAEVIRRLFDRTGIDPASIGYVEAHGTGTLLGDPVEVSALSEAFRTFTDRRGYCRLGSVKSNLGHLDTV
AGLAGLIKTALSLRQGEVPPTLHVTQVNPKLELTDSPFVIADRLAPWPSLPGPRRAAVSAFGLGGTNTHAI
LEHYPRDSRPRERSQRSNAVRAVAPFSARTLEALKDNLRALLDFLEDPASAEVALADITYTLQVGRVAMP
ERMVVTASTRDELVEGLRRGIATVGGAHVGTVVDTSPSVDADARAVAEAWATGDSIDWDSLHGDVKP
ARVSLPTYQFAKERYGLSPAHSVANSSKTHPDAGVPLFVPTWQPWSEGASNASLALRHLVVLCEPLDAL
GAEGASALASTLADRRIEVVRTSSPSARLDARFMAHASAVFERVKALLSERLTAPVTLQVLVPEERDALA
LSGLGSLLRSVSQENPLVRGQLIRVQGSVSASALVDVLVKSARAGDVTDSRYILAGQLSRCEWREARVAK
GDASRFWREDGVYVISGGTGALARLFVAEIGKRATRATVILVARASSAEAVDGGNGLRVRHLPVDVTQP
NDVNAFVATVLREHGRIDGVIHAAGIRRDNYLLNKPVAEMQAVLAPKVVGLVNLDHATRELPLDFFVTF
SSLAAFGNAGQSDYAAANGFMDGFAESRAALVNAGQRQGRTVSIRWPLWENGGMQLDSRSREVLMQR
TGMAALGDEAGLGAFYRALELGSPGVAVWTGEAQRFRELSVSVSPAPPPHQVALDAVVSITEKVETKLK
ALFSEVTRYEERRIDARQPMERYGIDSIIITQMNQALEGPYNALSKTLFFEYRTLAEVSGYLAEHRAEESA
KWVAAPGENSSSVIQEARPPRADATHRAPRADEPIAVIGMSGRYPGAENLTEFWERLSRGDDCITEIPPER
WSLDGFFYPDKKHAAARGMSYSKWGGFLGGFADFDPLFFNISPREATSMDPQERLFLQSCWEVLEDAG
YTRDSLAQRFGSAVGVFAGITKTGYELYGAELEGRDASVRPYTSFASVANRVSYLLDLKGPSMPVDTMC
SASLTAVHMACEALQRGACVMAIAGGVNLYVHPSSYVSLSGQQMLS

DNA sequence nucleotides 1-7178.

GTCGACCCGGCGAGGCTGACCCGGGCCTGGGAAGGACTGCTCGAACGGTATCCGCTGCTCGCTGGC
GCGATTCGCGTCGAAGGCACGGAGCCGGTCATCGTCCCCAGTGGGCAGGTCTCCGCCGAGGTCCAC
GAGGTTCCATCGGTCTCCGATTCAGCACTGGTGGCGACCCTGCGCGCCTCCGCGAAGGTGCCATTCG
ATCTCGCCTGTGGACCGCTCGCTCGGCTGCACCTGTACTCGCGGTCGGAGCACGAGCATGTCCTGCT
GCTGTGCTTCCACCACCTGGTGCTCGATGGGGCATCCGTGGCGCCCTTGCTCGACGCCCTCCGGGAG
CGTTACGCCGGGACCGAGGCGAAGGCGGGGCTGCTCGAGGTTCCGATCGTCGCTCCTTACCGCGCC
GCCGTGGAGTGGGAGCAGCTCGCCATTGGAGGCGATGAGGGACGGCGCCACCTCGACTACTGGCGG
CACGTGTTGGCCACGCCCGTTCCTCCGCCGTTGAATCTTCCAACGGACCGGCCTCGCTCCGCCACGG
GGCTGGACTCGGAGGGAGCAACGCACTCGCAGAGGGTGCCCACCGAGCAAGCATTGCGACTGCGCG
AGTTCGCTCGGGCACAGCAAGTGAGCCTGCCGACCGTCCTGCTCGGGCTCTACTACGCCTTGCTTCA
TCGGCACACGCGCCAGGACGACGTGGTGGTCGGCATCCCCACCATGGGGCGGCCCCGGGCGGAACT
GGCGACGGCGATTGGGTACTTCGTCAACGTGATGGCCGTGCGCGCGCGGGGCCTGGGGCAGCACTC
GTTCGGCTCGCTGCTGCGCCACCTCCACGACTCGGTCATCGATGGCCTGGAGCATGCCCACTATCCC
TTCCCGCGAGTGGTGAAGGACCTCCGGCTGTCGAATGGGCCCGAGGAGGCGCCTGGCTTCCAGACG
ATGTTCACCTTCCAGAGCCTGCAACTGACGAGCGCTCCGCCAAGGCCGGAGCCCAGGTCGGGCGGG
TTGCCGGAGCTTGAGCCGCTCGACTGCGTCCATCAGGAAGGCGCCTACCCGCTGGAGCTTGAAGTGG
TGGAGGGCGCCAAGGGCCTCACGCTGCATTTCAAGTACGACGCGCGGCTGTACGAGGCGGACACGG
TCGAACGGATGGCGCGTCAGTTGTTGCGCGCCGCGGACCAGGTCGCGGATGGGGTGGAGTCTCCGC
TGAGCGCACTGTCGTGGCTCGACGACGAAGAGCGCCGCACGCTTCTCCGCGACTGGAATGCCACGG
CCACGCCGTTCCTCGAGGACCTGGGCGTTCACGAGCTCTTCCAGCGGCAGGCCCGGGAGACCCCAG
ACGCCATGGCTGTGAGCTACGAGGGGCACTCGCTCAGCTATCAGGCGCTGGATACGCGGAGCCGCG
AGATTGCGGCGCACCTGAAGAGCTTCGGCGTCAAGCCTGGGGCGCTCGTGGGCATCTACCTGGACC
GGTCCGCGGAGCTGGTGGCGGCGATGCTGGGTGTGCTGTCCGCTGGCGCGGCCTACGTACCCCTGG
ACCCGGTGCACCCCGAGGACCGGCTGCGGTACATGCTGGAGGACAGTGGCGTGGTGGTCGTGCTGG
CCCGTCAGGCCTCGCGGGACAAGGTCGCCGCCATTGCCGGAGCCTCCTGCAAGGTGTGCGTGCTGG
AGGACGTCAAGGCTGGAGCCACGTCCGCGCCGGCGGGAACCTCACCGAACGGACTTGCCTACGTCA
TCTACACGTCCGGGAGCACGGGCCGGCCCAAGGGCGTGATGATTCCCCATCGCGGGGTGGTCAACT
TCCTCCTGTGCATGCGCAGGACGCTGGGCCTGAAGCGCACGGATTCGCTGTTGGCGGTCACGACGTA
CTGCTTCGACATCGCGGCGCTCGAGCTCCTGCTTCCGCTGTGTGCGGGGGCGCAGGTCATCATCGCG
TCGGCGGAGACGGTTCGGGATGCGCAGGCGTTGAAGCGGGCGCTGCGCACCCATCGGCCCACGTTG
ATGCAGGCGACGCCCGCGACCTGGACACTGTTGTTCCAGTCTGGCTGGGAGAACGCCGAGCGGGTT
CGAATCCTCTGCGGTGGAGAAGCGCTGCCGGAGTCGCTCAAGGCCCACTTCGTTCGCACCGCGAGC
GACGTGTGGAACATGTTCGGGCCCACCGAGACGACCATCTGGTCGACGATGGCGAAGGTCTCGGCC
TCGCGTCCGGTCACCATTGGAAAGCCGATCGACAACACGCAGGTCTACGTGCTGGACGACCGGATG
CAGCCGGTGCCCATCGGTGTGCCGGGCGAGCTGTGGATTGCGGGCGCGGGCGTGGCCTGCGGTTAC
CTCAACCGGCCGGCGCTGACCGCCGAGCGCTTCGTTTCCAATCCGTTCACGCCGGGCACGACGCTCT
ACCGGACGGGGGACCTGGCGCGCTGGCGCGCTGACGGTGAGGTTGAGTACCTGGGGCGGCTCGACC
ACCAGGTGAAGGTGCGCGGCTTCCGCATCGAGATGGGGGAGATTGAAGCGCAGTTGGCCGGGCATC
CCAGCGTGAAGAACTGTGCCGTGGTGGCCAAGGAGCTGAACGGCACCTCGCAGCTCGTCGCCTACT
GTCAGCCCGCGGGAACGAGCTTCGATGAGGAAGCCATCCGTGCACACCTGCGGAAGTTCCTCCCCG
ACTACATGGTCCCCGCGCACGTCTTCGCGGTGGATGCGATTCCGCTGTCGGGCAATGGCAAGGTGGA
CCGGGGCCAGCTGATGGCCAGGCCGGTGGTCACCCGGCGGAAGACATCCGCGGTCCATGCCCGTTC
GCCTGTTGAGGCCACCCTCGTCGAGCTGTGGAAGAACGTGCTCCAGGTCAACGAGGTGGGTGTCGA

GGATCGCTTCTTCGAAGTGGGGGGGGACTCCGTGCTGGCCGCCGTGCTGGTGGAGGAGATGAACCG
GCGCTTCGACACGCGGCTCGCCGTCACCGACCTGTTCAAGTACGTCAATATTCGCGACATGGCGCGC
CACATGGAGGGCGCGACGGCGCAAGCCCGTACTGGGGCCACCGAGCCGGCTCGCGAGGACACCGCG
TCGGAGCGTGACTACGAGGGCAGCCTGGCCGTCATCGGCATCTCCTGTCAGTTGCCCGGAGCCGCGG
ACCCCTGGCGCTTCTGGAAGAACCTGCGAGAGGGCAGGGACAGCGTGGTGGCGTACCGCCATGAGG
AACTGCGCGAGCTGGGCGTGCCCGAGGAGGTCCTCCGCGATTCCCGTTACGTCGCGGTCCGGTCGTC
CATCGAAGACAAGGAGTGCTTCGACCCGCATTTCTTCGGTCTGACGGCGCGGGACGCGTCCTTCATG
GACCCGCAGTTCCGACTGTTGCTGATGCACGCCTGGAAGGCAGTGGAAGACGCGGCGACGACGCCT
GAGCGCCTGGGACCGTGCGGCGTCTTCATGACGGCCAGCAACAGCTTCTATCACCAGGGCTCGCCGC
AATTTCCTGCGGACGGGCAGCCGGTCCTCCGCACCGCCGAAGAATACGTGCTGTGGGTGCTGGCCCA
GGCAGGCTCCATCCCGACGATGGTTTCSTACAAGCTCGGCTTGAAGGGGCCGAGCCTGTTCGTCCAC
ACCAACTGCTCGTCATCCCTGTCCGCGCTGTACGTGGCTCAGCAGGCCATCGCAGCGGGAGACTGCC
AGACGGCGCTGGTGGGGGCCGCCACGGTCTTCCCTTCGGCGAACTTGGGTTATCTGCACCAGCGGG
GGCTCAACTTCTCCAGCGCGGGGCGGGTCAAGGCCTTCGACGCCGCGGCGGACGGCATGATTGCCG
GTGAAGGTGTCGCCGTGCTGGTGGTGAAGGACGCCGCAGCGGCGGTGCGCGATGGCGACCCAATCT
ACTGCCTCGTGCGGAAGGTGGGGATCAACAACGACGGCCAGGACAAGGTGGGTTTATACGCCCCGA
GCGCCACCGGGCAGGCGGAGGTCATCCGGCGTCTGTTCGACCGGACCGGCATCGACCCTGCATCGA
TTGGCTACGTCGAGGCCCATGGCACCGGAACCTTGCTGGGTGACCCTGTCGAGGTCTCCGCGCTGAG
CGAAGCCTTCCGGACCTTCACCGACCGGCGCGGGTACTGCCGGCTGGGCTCGGTGAAGTCGAACCT
GGGCCATCTGGACACAGTGGCTGGACTGGCTGGGCTCATCAAGACGGCGCTGAGCCTGCGGCAGGG
CGAAGTTCCTCCGACGCTCCATGTGACCCAGGTGAATCCGAAGCTCGAGCTGACGGATTCGCCGTTC
GTCATCGCCGACCGTTTGGCGCCGTGGCCGTCCCTGCCGGGACCGAGGCGGGCGGCCGTGAGTGCG
TTCGGCCTTGGCGGGACGAATACCCACGCCATTCTCGAACACTACCCGCGCGACTCCCGCCCACGGG
AGAGGAGCCAGCGGTCGAACGCAGTCCGTGCGGTGGCTCCATTCTCGGCGCGCACCCTGGAGGCGT
TGAAGGACAACCTCCGCGCGCTGCTCGACTTCCTGGAGGACCCGGCGTCCGCGGAGGTGGCGCTCG
CGGACATCACCTACACGTTGCAGGTCGGCCGGGTCGCGATGCCTGAGCGGATGGTGGTGACTGCGT
CGACGCGCGACGAATTGGTGGAGGGACTGCGGCGAGGCATCGCGACGGTGGGCGGTGCCCACGTGG
GAACGGTGGTCGATACGTCACCCAGCGTGGATGCCGATGCTCGGGCAGTTGCGGAGGCGTGGGCGA
CGGGCGACTCGATTGACTGGGATTCGCTGCACGGTGACGTGAAGCCCGCCCGTGTCAGCCTGCCCAC
GTATCAGTTCGCGAAGGAGCGCTACGGGTTGTCGCCCGCGCACTCCGTGGCGAATTCCTCCAAGACG
CATCCTGACGCGGGTGTCCCGCTCTTCGTTCCGACCTGGCAGCCGTGGTCTGAGGGCGCGTCAAATG
CCTCGTTGGCGCTCCGGCACCTGGTGGTGTTCTGCGAGCCTCTTGATGCGCTGGGGGCTGAAGGTGC
CTCCGCGCTGGCGAGCACGCTCGCGGACAGGCGCATCGAAGTGGTCAGGACGTCCAGCCCAAGTGC
GCGGCTGGACGCGCGGTTCATGGCGCATGCCTCGGCGGTCTTCGAACGCGTCAAGGCGCTGCTGTCG
GAGCGTCTGACCGCTCCTGTGACATTGCAGGTGCTGGTGCCAGAGGAGCGGGATGCGCTGGCACTG
AGTGGCCTGGGGAGCCTGCTGCGTTCGGTGTCGCAGGAGAATCCGTTGGTCCGGGGGCAGCTCATC
CGCGTCCAGGGAAGCGTCTCCGCATCGGCGCTGGTGGACGTTCTGGTGAAGTCCGCGCGCGCCGGT
GACGTCACCGATTCGCGGTACCACGCGGGCCAGCTTTCTCGCTGTGAGTGGCGCGAGGCACGTGTCG
CCAAGGGGGACGCATCCCGCTTCTGGCGCGAAGACGGCGTCTATGTGATTTCAGGAGGAACCGGCG
CCCTGGCCCCGGCTGTTCGTCGCCGAAATCGGGAAGCGCGCGACGCGGGCCACCGTCATTCTGGTTGC
CCGCGCATCCTCGGCGGAGGCGGTGGACGGTGGGAACGGGCTGCGCGTGCGGCACCTTCCCGTGGA
TGTCACCCAACCGAACGACGTGAACGCCTTTGTCGCTACGGTGCTGCGCGAACACGGGCGCATCGAC
GGTGTCATCCATGCGGCGGGCATCCGCCGTGACAACTACCTGCTCAACAAGCCGGTGGCGGAAATG
CAGGCGGTGCTCGCGCCCAAGGTGGTGGGGCTCGTCAACCTGGACCACGCCACCCGCGAGCTGCCC
CTGGATTTCTTCGTCACGTTCTCGTCCCTGGCCGCGTTTGGAAACGCCGGTCAGTCGGACTACGCGG
CGGCCAATGGCTTCATGGACGGATTCGCGGAGTCCCGAGCGGCGCTCGTGAACGCCGGACAGCGGC
AGGGCCGGACGGTGTCCATCCGTTGGCCGCTCTGGGAGAACGGCGGGATGCAGCTCGACTCACGGA
GCCGTGAGGTCTTGATGCAGCGGACCGGGATGGCCGCGCTGGGAGACGAAGCGGGACTGGGGGCGT
TCTACCGGGCGCTGGAACTGGGCTCCCCTGGTGTCGCGGTGTGGACGGGGGAGGCCCAGAGGTTTC
GTGAACTCTCCGTGAGTGTTTCGCCCGCACCGCCTCCGCATCAGGTGGCGTTGGACGCCGTGGTGTC
CATCACCGAGAAGGTCGAGACGAAGCTGAAGGCGCTCTTCAGCGAGGTCACGCGATACGAAGAGCG
CCGCATCGATGCCCGCCAGCCGATGGAGCGCTATGGCATCGACTCCATCATCATCACGCAGATGAAC
CAAGCCCTCGAAGGGCCGTACAACGCCCTCTCGAAGACGCTGTTCTTCGAATACCGGACGCTCGCGG
AAGTCAGCGGGTATCTGGCCGAGCACCGCGCGGAAGAGAGCGCGAAGTGGGTGGCGGCACCTGGA
GAGAATTCGTCTTCCGTCATCCAGGAGGCCAGGCCGCCACGTGCGGATGCGACGCACCGGGCGCCT
CGCGCCGACGAGCCCATCGCCGTCATTGGCATGAGCGGCCGTTATCCCGGGGCGGAGAACCTGACG
GAGTTCTGGGAGCGCCTGAGCCGCGGTGACGACTGCATCACCGAGATTCCGCCAGAGCGCTGGTCG
TTGGACGGGTTCTTCTACCCGGACAAGAAGCACGCCGCCGCGCGGGGGATGAGCTACAGCAAGTGG
GGCGGCTTCCTCGGCGGCTTCGCTGACTTCGACCCGCTGTTCTTCAACATCTCGCCGCGTGAGGCGA
CGAGCATGGACCCGCAGGAGCGCTTGTTCCTGCAGAGCTGCTGGGAGGTCCTGGAGGACGCGGGGT
ACACCCGGGACAGCCTGGCCCAGCGCTTTGGCAGCGCGGTGGGCGTTTTCGCGGGAATCACGAAGA
CGGGCTACGAACTCTACGGCGCGGAGCTGGAAGGACGAGATGCCTCGGTCCGGCCCTATACGTCGT
TTGCGTCTGTTGCCAACCGCGTCTCGTATCTGCTCGACCTGAAGGGGCCGAGCATGCCCGTGGACAC
CATGTGCTCGGCCTCGCTGACAGCCGTCCACATGGCTTGCGAGGCGCTGCAACGAGGCGCCTGCGTC

2290.00075

ATGGCCATCGCGGGTGGAGTGAATCTCTACGTCCACCCGTCGAGCTACGTCAGCCTGTCCGGGCAGC
AGATGCTGTCGAC

REGION 2

TaR1 - Surface layer protein

From nucleotide 2955 to 601, size(aa): 785.

MKVVNKLLEKLPDVVAGKVPDVKLQDQDIKVPLAQGTFTEEKILPPKLAMHGFTLSFEATGEASIRNFNS
LGDVDENGIIGEPSPESAEPGPRPQLLLGSDIGWMRYQVSARVKAAVSASLSFLASENQTELSVTLSDYRA
HPLGQNMREAVRSDLSELRLMQATDLAKLTTGDAVAWHVRGALHTRLELNWADIFPTNLNRLGFLRGN
ELLALKTSAKAGLSARVSLTDDYQLSFSRPRAGRIQVAVRKVKSHEQALSAGLGITVELLDPATVKAQLG
QLLEALLGPVLRDLVKKGTTAVEIMDGLVDKASKAKLDDNQKKVLGLVLERLGIDPQLADPANLPQAW
ADFKARVAESLENAVRTQVAEGFEYEYLRLSETSTLLEVVVEDVTAMRFHESLLKGNLVELLKWMKSLP
AQQSEFELRNYLHATTLTRQQAIGFSLGLGSFELLKAKNVSKQSWVTQENFQGARRMAFLGRRGYEDKL
LGTRGQWVVDLKADMTRFSPTPVASDFGYGLHLMLWGRQKKLSRKDLQQAVDDAVVWGVLDAKDA
ATVISTMQEDMGKHPIETRLELKMADDSFRALVPRIQTLELSRFSRALARALPWSEQLPRASAEFRRAVY
APIWEAYLREVQEQGSLMLNDLSPSRAAQIAKWYFQKDPTVRDLGKDLQLIESEWRPGGGNFSFAEVIS
KNPNTLMRCRNFVSGMVRLRRAIDERKAPDELRTVFGELEGMWTTGFHLRAAGSLLSDLAQSTPLGLAG
VERTLTVRVADSEEQLVFSTARSTGAA

TaR2 - two component system, response regulator

From nucleotide 3116 to 4702, size(aa): 529.

MPSGCYGAASAFVLPPLPAMPQAPSDVSQVLLPFGGLVGREVDLDAFLQTLMDRIAITLQADRGTLWLL
DPARRELFSRAAHLPEVSQIRVKLGQGVAGTVAKAGHAINVPDPRGEQRFFADIDRMTGYRTTSLLAVPL
RDGDGALYGVLQVLNRRGEDRFTDEDTQRLTAIASQVSTALQSTSLYQELQRAKEQPQVPVGYFFNRIIG
ESPQLQAIYRLVRKAAPTDATVLLRGESGSGKELFARAVHVNGPRRDQPFIKVDCAALPATLIENELFGH
ERGAFTGADHRVPGKFEAASGGTVFIDEIGELPLPVQGKLLRVIQDREFERVGGTQAVKVDVRIVAATHR
DLARMVAEGRFREDLYYRIKVVEVVLPPLRERGAEDIERLARHFVAAVARRHRLTPPRLSAAAVERLKR
YRWPGNVRELENCIESAVVLCEGEILEEHLPLPDVDRAALPPPAAAQGVNAPTAPAPLDAGLLPLAEVER
RHILRVLDAVKGNRTAAARVLAIGRNTLARKLKEYGLGDEP

TaR3 - two component system, kinase sensor.

From nucleotide 5595 to 4720, size(aa): 292

MRASQAEAPHSRRLTMEVRFHGVRGSIAVSGSRIGGNTACVEVTSQGHRLILDAGTGIRALGEIMMREG
APQEAILFFSHLHWDHVQGFPFFTPAWLPTSELTLYGPGANGAQALQSELAAQMQPLHFPVPLSTMRSR
MDFRSALHARPVEVGPFRVTPIDVPHPQGCLAYRLEADGHSFVYATDVEVRVQELAPEVGRLFEGADVL
CLDAQYTPDEYEGRKGVAKKGWGHSTMMDAAGVAGLVGARRLCLFHHDPAHGDDMLEDMAEQARA
LFPVCEPAREGQRLVLGRAA

TaA - NUS-G like transcription antitermination.

From nucleotide 6290 to 6793, size(aa): 168

MPGPRCAENDWVALLVRVNHEKVAAAQLGKHGYEFFLPTYTPPKSSGVKAKLPLFPGYLFCRYQPLNP
YRIVRAPGVIRLLGGDAGPEAVPAQELEAIRRVADSGVSSNPCDYLRVGQRVRIIEGPLTGLEGSLVTSKS
QLRFIVSVGLLQRSVSVEVSAEQLEPITD

TaB - acyl carrier protein (ACP).

From nucleotide 6870 to 7106, size(aa): 79

MDKRIIFDIVTSSVREVVPELESHPFEPEDDLVGLGANSLDRAEIVNLTLEKLALNIPRVELIDAKTIGGLV
DVLHARL

TaC - beta-ketoacyl [ACP] synthase III (KAS III, FabH)

From nucleotide 7119 to 8378, size(aa): 420

MGPVGIEAMNAYCGIARLDVLQLATHRGLDTSRFANLLMEEKTVPLPYEDPVTYGVNAARPILDQLTAA
ERDSIELLVACTESSFDFGKAMSTYLHQHLGLSRNCRLIELKSACYSGVAGLQMAVNFILSGVSPGAKAL
VVASDLSRFSIAEGGDASTEDWSFAEPSSGAGAVAMLVSDTPRVFRVDVGANGYYGYEVMDTCRPVAD
SEAGDADLSLLSYLDCCENAFREYTRRVPAANYAESFGYLAFHIPFGGMVKGAHRTMMRKFSGKNRGD
IEADFQRRVAPGLTYCQRVGNIMGATMALSLLGTIDHGDFATAKRIGCFSYGSGCSSEFFSGVVTEEGQQ
RQRALGLGEALGRRQQLSMPDYDALLKGNGLVRFGTRNAELDFGVVGSIRPGGWGRPLLFLSAIRDFHR
DYQWIS

TaD - membrane associated protein

From nucleotide 8404 to 9378, size(aa): 325

MSSVATAVPLTARDSAVSRRLRITPSMCGQTSLFAGQIGDWAWDTVSRLCGTDVLTATNASGAPTYLAF
YYFRIRGTPALHPGALRFGDTLDVTSKAYNFGSESVLTVHRICKTAEGGAPEADAFGHEELYEQPQPGRI
YAETFNRWITRSDGKSNESLIKSSPVGFQYAHLPLLPDEYSPRRAYGDARARGTFHDVDSAEYRLTVDRF
PLRYAVDVIRDVNGVGLIYFASYFSMVDWAIWQLARHQGRSEQAFLSRVVLDQQLCFLGNAALDTTFDI
DVQHWERVGGGEELFNVKMREGAQGRDIAVATVKVRFDAASEGGRRG

TaE - acyl carrier protein (ACP).

From nucleotide 9386 to 9364, size(aa): 82

MTDEQIRGVVHQSIVRVLPRVRSNEIAGHLNLRELGADSVDRVEILTSILDSLRLQKTPLAKFADIRNIDAL
VAFLAGEVAGG

TaF - beta-ketoacyl [ACP] synthase III (KAS III, FabH)

From nucleotide 9757 to 10878, size(aa): 374

MMQERGVALPFEDPVTNAVNAARPILDAMSPEARERIELLVTSSESGVDFSKSISSYAHEHLGLSRHCRFL
EVKQACYAATGALQLALGYIASGVSPGAKALVIATDVTLVDESGLYSEPAMGTGGVAVLLGDEPRVMK
MDLGAFGNYSYDVFDTARPSPEIDIGDVDRSLFTYLDCLKHSFAAYGRRVDGVDFVSTFDYLAMHTPFA
GLVKAGHRKMMRELTPCDVDEIEADFGRRVKPSLQYPSLVGNLCSGSVYLSLCSIIDTIKPERSARVGMF
SYGSGCSSEFFSGVIGPESVSALAGLDIGGHLRGRRQLTFDQYVELLKENLRCLVPTKNRDVDVERYLPL
VTRTASRPRMLALRRVVDYHRQYEWV

TaG - signal peptidase II (LSPA)

From nucleotide 10909 to 11421, size(aa): 171

MNTPSLTNWPARLGYLLAVGGAWFAADQVTKQMARDGAKRPVAVFDSWWHFHYVENRAGAFGLFSS
FGEEWRMPFFYVVGAICIVLLIGYYFYTPPTMKLQRWSLATMIGGALGNYVDRVRLRYVVDFVSWHVG
DRFYWPSFNIADTAVVVGAALMILESFREPRQQLSPG

TaH - cytochrome P450 hydroxylase (cP450)

From nucleotide 11473 to 12897, size(aa): 475

MGTSEPVEPDHALSKPPPVAPVGAQALPRGPAMPGIAQLMMLFLRPTEFLDRCAARYGDTFTLKIPGTPP
FIQTSDPALIEVIFKGDPDLFLGGKANNGLKPVVGENSLLVLDGKRHRRDRKLIMPTFLGERMHAYGSVI
RDIVNAALDRWPVGKPFAVHEETQQIMLEVILRVIFGLEDARTIAQFRHHVHQVLKLALFLFPNGEGKPA
AEGFARAVGKAFPSLDVFASLKAIDDIIYQEIQDRRSQDISGRQDVLSLMMQSHYDDGSVMTPQELRDEL
MTLLMAGHETSATIAAWCVYHLCRHPDAMGKLREEIAAHTVDGVLPLAKINELKFLDAVVKETMRITP
VFSLVARVLKEPQTIGGTTYPANVVLSPNIYGTHHRADLWGDPKVFRPERFLEERVNPFHYFPFGGGIRK
CIGTSFAYYEMKIFVSETVRRMRFDTRPGYHAKVVRRSNTLAPSQGVPIIVESRLPS

TaI - malonyl CoA [ACP] transacylase (MCT, FabD)

From nucleotide 12938 to 13891, size(aa): 318

MVDSVSKQARRKVFLFSGQGTQSYFMAKELFDTQTGFKRQLLELDEQFKQRLGHSILEIIYDARAARLD
PLDDVLVSFPAIFMIEHALARLLIDRGIQPDAVVGASMGEVAAAAIAGAISVDAAVALVAAQAQLFARTA
PRGGMLAVLHELEACRGFTSVARDGEVAAINYPSNFVLAADEAGLGRIQQELSQRSVAFHRLPVRYPFHS
SHLDPLREEYRSRVRADSLTWPRIPMYSCTTANRVHDLRSDHFWNVVRAPIQLYDTVLQLEGQGGCDFI
DVGPAASFATIIKRILARDSTSRLFPLLSPSPASTGSSMG

TaJ - malonyl CoA [ACP] transacylase (MCT, FabD)

From nucleotide 13909 to 14898, size(aa): 330

MTEAPAPRAPAQVPPPPSSPWALHTRGAASAPVNARKAALFPGQGSQERGMGAALFDEFPDLTDIADAI
LGYSIKRLCLEDPGKELAQTQFTQPALYVVNALSYLKRLREGAEQPAFVAGHSLGEYNALLVAGAFDFE
TGLRLVKRRGELMSGASGGTMAAVVGCDAVAVEQVLRDRQLTSLDIANINSPDQIVVSGPAQDIERARQ
CFVDRGARYVPLNVRAPFHSRYMQPAASEFERFLSQFQYAPLRCVVISNVTGRPYAHDNVVQGLALQLR
SPVQWTATVRYLLEQGVEDFEELGPGRVLTRLITANKRGAPAPATAAPAKWANA

TaK - 3-oxoacyl [ACP] synthase (KAS I, FabB)

From nucleotide 14963 to 16213, size(aa): 417

MSTSPVQELVVSGFGVTSAIGQGAASFTSALLEGAARFRVMERPGRQHQANGQTTAHLGAEIASLAVPE
GVTPQLWRSATFSGQAALVTVHEAWNAARLQAVPGHRIGLVVGGTNVQQRDLVLMQDAYRERVPFLR
AAYGSTFMDTDLVGLCTQQFAIHGMSFTVGGASASGLLAVIQAAEAVLSRKVDVCIAVGALMDVSYWE
CQGLRAMGAMGTDRFAREPERACRPFDRESDGFIFGEACGAVVVESAEHARRRGVTPRGILSGWAMQL
DASRGPLSSIERESQVIGAALRHADLAPERVDYVNPHGSGSRQGDAIELGALKACGLTHARVNTTKSITG
HGLSSAGAVGLIATLVQLEQGRLHPSLNLVDPIDSSFRWVGATAEAQSLQNALVLAYGFGGINTAVAVR
RSATES

TaL - enoyl CoA hydratase.

From nucleotide 16224 to 17009, size(aa): 262

MQAASPPHRDYQTLRVRFEAQTCFLQLHRPDADNTISRTLIDECQQVLTLCEEHATTVVLEGLPHVFCM
GADFRAIHDRVDDGRREQGNAEQLYRLWLQLATGPYVTVAHVQGKANAGGLGFVSACDIVLAKAEVQ
FSLSELLFGLFPACVMPFLARRIGIQRAHYLTLMTRPIDAAQALSWGLADAVDADSEKLLRLHLRRLRCLS
KPAVTQYKKYASELGGQLLAAMPRAISANEAMFSDRATLEAIHRYVETGRLPWES

TaM - enoyl CoA hydratase.

From nucleotide 17000 to 17767, size(aa): 256

MGIMTEGTPMAPVVTLHEVEEGVAQITLVDRENKNMFSEQLVRELITVFGKVNGNERYRAVVLTGYDT
YFALGGTKAGLLSICDGIGSFNVTNFYSLALECDIPVISAMQGHGVGGGFAMGLFADFVVLSRESVYTTN
FMRYGFTPGMGATYIVPKRLGYSLGHELLLNARNYRGADLEKRGVPFPVLPRKEVLPHAYEIARDLAAK
PRLSLVTLKRHLVRDIRRELPDVIEREIEMHGITFHHDDVRRRIEQLFL

TaN - O-methyltransferase (fragment).

From nucleotide 17782 to 19053, size(aa): 423

MLNLINNHAHGYVVTPVVLACNDAGLFELLRQGPKDFDRLAEALRANRGHLRVAMRMFESLGWVRRD
ADDVYAVTAAAAAHRSFPREAQSLFALPMDRYLRGEDGLSLAPWFERSRASWDTDDTLVRELLDGAIIT
PLMLALEQRGGLKEARRLSDLWSGGDGRDTCVPEAVQHELAGFFSAQKWTREDAVDAELTPKGAFIFE
RALLFAIVGSYRPMLASMPQLLFGDCDQVFGRDEAGHELHLDRTLNVIGSGHQHRKYFAELEKLIITVFD
AENLSAQPRYIADMGCGDGTLLKRVYETVLRHTRRGRALDRFPLTLIAADFNEKALEAAGRTLAGLEHV
ALRADVARPDRLIEDLRARGLAEPENTLHIRSFLDHDRPYQPPADRAGLHARIPFDSVFVGKAGQEVVPA
EVFHSLVEHLE

DNA sequence 1-19053

GTCGACGTTGACGTCGCCCGGTGGCGTGCCGTGTGTCTTCTTCGACGCGGAGGTGCGCGAGGTGGCG
GCGGACGGCCGGCGCGGGCCGCTGTTGTCGCGTGAGCGCGCGTATGCGCCGGTACTGGCGCTGCGT
GGCCAGCGCCTCCATGCTTCGGTGTCCTTTTCGCCCGCGTCGCTGATGGCTCCGGTGGAGGTGCGCC
GGTGCAAGGCCCTGCCAGGCACGGTGCCCGCGTCCTGGTATCAGACGGCGCACCCGGAGGCCCTGT
CCTGGGAGCGCGTGGGCGCGGTGGGCGAATCCTGCCTCGTGGTGGGTGAACTCCGGAGGGGCCCTG
TCGAGGGCAGCTACGCCCTGGTCGGTCGCGAGGGCGGCCCCGCGATGTTGGTGCTGGGACCCCAGG
CTCCGGCCACCTGTGGGACGCTGGCGCGCCGGGCCTGGCGGCACTTCGCGGCGGCCGGGGTGCTGT
CCATGGCCGCGGCCGTCGTCCTGTCAGGGGCGCTGTGAGACGCGCGGCGGGGGCCGTACCGCCGCG
CCAGAAACGTGATGCGCCGCCAGGCCTCGCGGTCCGGGCACTGACGCCCGGGCCGCTCGGGACTCG
CTCAGGCGGCTCCGGTGCTTCGCGCGGTGGAGAACACGAGCTGTTCCTCGCTGTCCGCCACCCGCAC
GGTGAGGGTCCGCTCCACGCCGGCGAGGCCCAGCGGCGTGGACTGCGCCAGGTCCGAGAGCAGGGA
GCCCGCAGCGCGCAGGTGGAAGCCGGTGGTCCACATGCCCTCCAGCTCGCCGAACACGGTGCGCAG
CTCGTCCGGGGCCTTGCGTTCGTCGATGGCGCGGCGCAGGCGCACCATGCCGCTCACGAAGTTCCTG
CACCGCATGAGCGTGTTGGGGTTCTTGGAGATGACCTCCGCGAAGCTGAAGTTGCCGCCACCCGGGC
GCCACTCGCTTTCGATGAGCTGCAGGTCCTTGCCAAGGTCGCGCACCGTGGGGTCCTTCTGGAAGTA
CCACTTGGCGATCTGCGCGGCGCGGCTGGGTGACAAGTCATTCAGCATGAGGCTGCCTTGCTCCTGC
ACCTCGCGGAGGTAGGCCTCCCAGATGGGGGCGTAGACCGCGCGCCGGAACTCGGCGGAGGCGCGG
GGAAGCTGCTCGCTCCAGGGCAGCGCGCGGGCCAGGGCGCGTGAGAAGCGGGACAGCTCGAGCGTC
TGGATGCGGGGCACCAGGGCGCGGAACGAGTCATCCGCCATCTTCAGCTCGAGCCGCGTTTCGATG
GGGTGCTTGCCCATGTCCTCCTGCATGGTGCTGATGACGGTGGCCGCGTCCTTCGCGTCCAGCACGC
CCCAGACGACGGCGTCATCCACCGCCTGCTGCAGGTCCTTGCGCGACAGCTTCTTCTGCCGTCCCCA
CAGCATCAGGTGCAGGCCGTAGCCGAAGTCGGAGGCCACGGGGGGTGGGAGAGAAGCGCGTCATGTC
CGCCTTCAGGTCCACCACCCACTGGCCGCGGGTGCCCAGCAGCTTGTCCTCGTAGCCCCGGCGTCCG
AGGAACGCCATGCGCCGGGCGCCCTGGAAGTTCTCCTGCGTCACCCAGGACTGCTTGCTGACGTTCT
TCGCCTTGAGCAGCTCGAACGAGCCCAGCCCCAGTGAGAAGCCGATGGCCTGCTGGCGCGTGAGCG
TGGTGGCGTGCAGGTAGTTGCGCAGCTCGAACTCGCTCTGCTGGGCGGGGAGGCTCTTCATCCACTT
CAGCAGCTCCACCAGGTTGCCCTTGAGCAGGGACTCGTGGAAGCGCATCGCGGTGACGTCCTCCACG
ACGACCTCCAGCAGCGTGGAGGTCTCCGACAGGCGCAGGTATTCGTACTCGAAGCCCTCGGCGACCT
GCGTGCGGACGGCGTTCTCCAGCGACTCTGCGACGCGGGCCTTGAAGTCGGCCCAGGCCTGCGGAA
GGTTGGCCGGGTCCGCAAGCTGCGGGTCGATGCCAAGGCGCTCCAGCACCAGGCCCAGCACCTTCTT
CTGATTGTCGTCCAGCTTCGCCTTGCTGGCCTTGTCCACCAGGCCGTCCATGATTTCCACCGCGGTGG
TGCCCTTCTTGACGAGGTCGCGAAGGACGGGCCCCAGCAGCCGCTTCCAGCAACTGGCCCAGTTGGG-
CCTTCACCGTCGCCGGGTCCAGCAGCTCCACGGTGATGCCCAGGCCGGCGGAGAGCGCCTGCTCATG
GGACTTCACCTTGCGCACGGCGACCTGGATGCGGCCGGCACGGGGACGGGAGAAGCTGAGCTGGTA
GTCGTCGGTGAGGGACACCCGGGCGGACAGGCCCGCCTTGGCGCTGGTCTTCAACGCGAGCAGCTC

2290.00075

GTTGCCGCGCAGGAAGCCCAGGCGGTTGAGGTTGGTGGGGAAGATGTCCGCCCAGTTGAGCTCCAG
CCGTGTGTGGAGCGCGCCGCGGACATGCCACGCCACCGCGTCCCCCGTGGTCAGCTTGGCCAGGTCG
GTGGCCTGCATCAGCCGCAGCTCGGACAGGTCGGAGCGCACGGCCTCACGCATGTTCTGGCCCAGC
GGATGCGCGCGGTAGTCGCTGAGCGTGACGGACAGCTCCGTCTGGTTCTCGGAGGCGAGGAAGGAC
AGGCTGGCGCTCACGGCGGCCTTCACGCGCGCGGACACCTGGTAGCGCATCCACCCGATGTCACTGC
CCAGCAGCAGTTGGGGCCGGGGCCCTGGCTCGGCGCTCTCCGGGCTCGGCTCGCCGATGATGCCGTT
TTCGTCCACGTCGCCCAGCGAGTTGAAGTTCCGGATGGACGCTTCGCCGGTGGCTTCGAAGGAGAG
GGTGAAGCCGTGCATGGCGAGCTTGGGCGGAAGGATTTTCTCTTCCGTGAAGGTCCCCTGGGCCAGC
GGCACCTTGATGTCCTGGTCCTGCAGCTTCACGTCGGGCACCTTGCCCGCCACGACGTCGGGAAGCT
TCTCCAGCAGCTTGTTGACCACTTTCATGCGCGTCCCCCTGGGCTGAAGCCTCCTGCACGTGGGCCG
GAGGTCTCTTCGTCGTACGCCGTTGCCCAGCTCGGAACAAGGCGGATACCAGAAAAGACCGGTGGT
CAGCGGACAGATGCCCTGGAGGGTGGGGTGGGAGCCGCCCCCGCGCGGTGCGTCAGGGCTCGTCGC
CCAATCCGTACTCCTTGAGTTTCCGCGCGAGCGTGTTGCGGCCAATCGCCAGCACGCGGGCCGCGGC
GGTGCGGTTGCCCTTCACGGCGTCCAGCACGCGCAGGATGTGGCGGCGTTCGACCTCCGCCAGTGGC
AGCAGGCCCGCATCCAGGGGCGCAGGCGCAGTCGGCGCGTTGACACCCTGAGCGGCTGCGGGAGGC
GGCAGGGCGGCCCGGTCCACATCGGGCAGGGGCAGGTGCTCTTCGAGAATCTCCCCTTCACAGAGC
ACCACGGCGCTCTCGATACAGTTCTCCAGCTCCCGCACGTTTCCGGGCCAGCGGTAGCGCTTGAGGC
GCTCCACCGCGGCGGCGCTGAGGCGGGGCGGCGTCAGCCGGTGCCTCCGGGCGACGGCGGCGACGA
AGTGGCGGGCGAGCCGCTCGATGTCCTCCGCGCCGCGCTCCCGCAGCGGCGGCAGCACCACCTCGA
CCACCTTGATGCGGTAGTAGAGGTCCTCGCGGAAGCGGCCCTCGGCCACCATGCGGGCCAGGTCCC
GATGGGTGGCCGCGACGATGCGCACGTCCACCTTCACGGCCTGGGTGCCTCCCACGCGCTCGAACTC
GCGATCCTGGATGACCCGCAGCAACTTGCCCTGCACCGGCAGGGGCAGCTCGCCAATCTCGTCGATG
AACACGGTGCCGCCGCTGGCCGCTTCGAACTTGCCGGGCACGCGGTGGTCCGCGCCGGTGAAGGCG
CCGCGTTCGTGGCCGAAGAGCTCGTTCTCGATGAGCGTGGCGGGCAGCGCCGCGCAGTCCACCTTGA
TGAAGGGCTGGTCCCTGCGGGGACCATTCACGTGGACGGCACGGGCGAACAGCTCCTTGCCGCTGC
CACTCTCGCCGCGCAGCAGCACCGTCGCATCGGTGGGCGCGGCCTTGCGCACCAGTCGGTAGATGG
CCTGGAGCTGCGGGGACTCGCCGATGATGCGGTTGAAGAAGTAGCCCACCGGTACCTGGGGCTGCT
CCTTCGCGCGCTGGAGCTCTTGATAGAGGCTGGTGCTCTGGAGGGCGGTGCTCACCTGCGAGGCGAT
GGCGGTGAGCCGCTGCGTGTCCTCGTCGGTGAAGCGGTCCTCGCCGCGGCGGTTGAGGACCTGGAG
CACGCCGTAGAGGGCGCCGTCCCCGTCGCGCAGTGGCACGGCGAGCAGGCTGGTGGTGCGGTAGCC
CGTCATCCGGTCGATGTCCGCGAAGAAGCGCTGCTCGCCGCGCGGGTCCGGCACGTTGATGGCGTG
CCCCGCCTTGGCGACGGTGCCGGGCGACGCCCTGGCCCAGCTTGACGCGAATCTGGGACACCTCGGGC
AGGTGCGCGGCGCGGCTGAACAGCTCGCGGCGGGCCGGGTCCAGCAGCCAGAGCGTGCCGCGGTCC
GCTTGCAGGGTGATGGCGATGCGGTCCATCAGCGTCTGGAGGAACGCGTCGAGGTCCACCTCCCTG
CCGACGAGTCCTCCGAAGGGGAGGAGGACCTGGCAGACGTCCGAGGGGGCTTGGGGCATGGCGGG
CAACGGCGGCAGGACGAAGGCGGAGGCCGCACCATAACATCCAGAGGGCATGGGACTGCCCCCTCT
CAGGCCGCGCGGCCCAGCACCAGCCGCTGGCCTTCGCGTGCGGGCTCGCACACGGGGAAGAGGGCG
CGGGCCTGCTCCGCCATGTCCTCGAGCATGTCGTCGCCGTGCGCCGGGTCATGGTGGAACAGGCACA
GCCGGCGCGCCCCCACCAGCCCGGCCACGCCCGCGGCATCCATCATGGTGGAGTGGCCCCAGCCCTT
CTTCGCCACGCCCTTGCGGCCCTCGTATTCGTCCGGCGTGTACTGCGCATCCAGGCACAGGACGTCC
GCGCCCTCGAAGAGGCGGCCCACCTCCGGCGCGAGCTCCTGCACCCGCACCTCCACGTCCGTGGCGT
AGACGAACGAATGGCCATCCGCCTCCAGGCGGTACGCCAGGCACCCCTGCGGGTGCGGCACGTCGA
TGGGCGTGACGCGGAAGGGGCCCACCTCCACGGGTCGGGCATGCAACGCCGAGCGGAAGTCCATCC
GCGAGCGCATGGTGCTCAGCGGCACCGGAAAATGAAGCGGCTGCATCTGCGCGGCCAACTCGGACT
GGAGCGCCTGGGCCCCATTCGCGCCCGGACCGTAGAGCGTCAGCTCGGACGTGGGCAGCCAGGCCG
GCGTGAAGAAGGGGAAGCCCTGCACGTGGTCCCAATGCAGATGCGAGAAGAAGAGCGTGGCCTCCT
GGGGCGCGCCCTCGCGCATCATGATTTCGCCCAGTGCGCGGATGCCCGTCCCCGCATCCAGGATGAG
GCGGTGGCCCTGGCTGGTCACCTCCACGCAGGCCGTGTTGCCACCAATGCGCGAGCCCGACACCGCG
ATGCTCCCCCGAACGCCATGAAACCGGACTTCCATCGTAAGTCTCCTTGAATGGGGGGCCTCCGCCT
GGGACGCCCTCATGCCCGGAGCCTCAGAGCACGGGGTGTGCCATTCCCAAATGCCCGGAATCAGGA
GCGCGGGCCTCGGGCTCGTCCACCGGTGCTCCAGAATGGATCGCGCTCGCCTGGTGCGGGCGATCCA
AAGCGGTGCAGGTCGCCCGCAGGACGGGGCGGCGGGCACGTCTTCCAACGTCCCACGGCAGTCCTG
TCTTCAGATCTCTCCCGATGCGGGAAGGCGTCCAGGAGGTTGCACCCGGCATCGAGCGGGGCTGTGT
GTTTCAAGTCTTGTCGGAGCCTCGGACACAACCGTCTGGGTTCTGGGAATGCGCCGGCTTCCGTTCA
CTCCAGAGTGATTCAATGGCTCTCGAGTGCAGGTTTAGCAATCCTCGGGCCGTAACCACGCCGTTGA
AGGCAGTCACGCTCTCGTCACGCTTGGGGTGTTTCCAGCTTCAACGGTGTTTATCCTTCAGGGCGGT
TTGCTTGACACGCTGCCTCATGGAAGCGTATGCAAAACAATGAAAACGGTGTCGTTGCCGAGCCTTA
GGGCCTCCAGAACGCCATCCTCGCGGACCCAGGCAGCCGGAATTTGAGACGGGGCTGTCAGCGGTT
TGAACGCAAGGATGCGGCGGGGGTTGTGGCGGCAGCCCGACCAGAATTCGGTTGGTGTGCCAGTTA
TTGTCAGATTCTGAGAAATAGCAGGCTGGGGGGAAGTTGCAATGCCTGGGCCGCGGTGTGCTGAGA
ACGATTGGGTTGCATTGCTCGTCCGCGTCAATCACGAGAAAGTGGCTGCCGCTCAGTTGGGGAAACA
CGGCTACGAGTTCTTCCTGCCGACGTACACGCCTCCCAAGTCCTCGGGTGTGAAGGCGAAGCTTCCG
CTCTTCCCCGGGTACCTTTTCTGTCGTTACCAGCCGCTCAATCCGTACCGCATCGTCCGGGCGCCCGG
GGTCATCCGGCTGCTCGGAGGTGACGCGGGGCCGGAAGCCGTGCCCGCACAGGAATTGGAGGCCAT

2290.00075

CCGCCGGGTCGCGGATTCGGGTGTCTCTTCCAATCCCTGTGACTATCTGCGGGTGGGGCAGCGCGTG
CGCATCATCGAAGGGCCCCTGACAGGTCTGGAAGGAAGTCTGGTGACGAGCAAGAGCCAACTCCGG
TTCATTGTCTCCGTGGGGCTGCTACAGCGCTCGGTGTCCGTGGAGGTGAGCGCCGAGCAACTGGAAC
CGATCACCGACTGATTCCGCGGACATCCCCTTCCATTCCTTCATCACCCCGACCCGCAGCAAGGCTTC
AGGGACCGTGAGTCGTTCCATGGACAAGAGAATTATTTTCGACATCGTCACCAGCAGTGTTCGGGAG
GTGGTACCCGAACTCGAATCACATCCGTTCGAGCCGGAGGATGACCTGGTCGGACTGGGCGCGAAC
TCGCTCGACCGCGCCGAAATCGTCAACCTCACGCTGGAGAAGCTGGCGCTCAACATCCCCCGGGTCG
AGCTGATTGACGCGAAGACCATTGGCGGGCTGGTGGACGTCCTTCACGCGAGGCTGTGAGGCGAAG
CCATGGGGCCGGTCGGGATTGAAGCCATGAATGCCTACTGTGGCATCGCCAGGTTGGATGTGTTGCA
GCTGGCGACCCACCGTGGCCTGGACACCTCCCGCTTCGCGAACCTGCTCATGGAGGAGAAGACCGTC
CCGCTCCCCTATGAGGACCCTGTCACCTACGGCGTGAATGCCGCCCGGCCCATCCTGGACCAGTTGA
CCGCGGCGGAACGGGACAGCATCGAGCTGCTGGTGGCTTGCACGGAGTCCTCGTTCGACTTCGGCA
AGGCCATGAGCACCTACCTGCACCAGCACCTGGGGCTGAGCCGCAACTGCCGGCTCATCGAGCTCA
AGAGCGCCTGCTACTCCGGGGTCGCCGGGCTGCAGATGGCCGTCAACTTCATCCTGTCCGGCGTGTC
GCCGGGGGCCAAGGCCCTGGTGGTGGCCTCCGACCTGTCGCGCTTCTCCATCGCCGAAGGGGGAGA
TGCCTCCACGGAGGACTGGTCCTTCGCGGAGCCGAGCTCGGGTGCGGGCGCGGTGGCCATGCTGGT
GAGCGACACGCCCCGGGTGTTCCGCGTCGACGTGGGGGCGAACGGCTACTACGGCTACGAGGTGAT
GGATACCTGCCGCCCGGTGGCGGACAGCGAAGCGGGAGACGCGGACCTGTCGCTCCTCTCGTACCT
GGACTGCTGTGAGAACGCCTTCCGGGAGTACACCCGCCGCGTCCCCGCGGCGAACTACGCGGAGAG
CTTCGGCTACCTCGCCTTCCACACGCCGTTTGGCGGCATGGTGAAGGGCGCCCACCGCACGATGATG
CGCAAGTTCTCCGGCAAGAACCGCGGGGACATCGAAGCGGACTTCCAGCGGCGAGTGGCCCCCGGG
CTGACCTACTGCCAGCGCGTGGGGAACATCATGGGCGCGACGATGGCGCTCTCGCTCCTCGGGACC
ATCGACCACGGCGACTTCGCCACCGCGAAGCGGATTGGCTGCTTCTCGTATGGCTCGGGGTGCAGCT
CGGAGTTCTTCAGCGGCGTGGTGACGGAGGAAGGGCAGCAGCGGCAGCGCGCCCTGGGGCTGGGA
GAAGCGCTGGGGCGCCGGCAGCAGCTCTCCATGCCGGATTACGACGCGCTGCTGAAGGGGAACGGC
CTGGTGCGCTTCGGGACCCGGAACGCCGAGCTGGATTTCGGTGTCGTCGGCAGCATCCGGCCGGGC
GGGTGGGGCAGGCCCTTGCTCTTCTTGTCGGCGATTCGTGACTTCCATCGCGACTACCAATGGATTT
CCTAGCCTCGGGGCTTCGAGCAAAGCCATGTCCAGCGTAGCGACGGCCGTCCCCCTGACGGCCCGTG
ACAGCGCGGTGAGCCGCCGGCTGCGAATCACCCCCAGCATGTGCGGCCAGACGTCCTTGTTCGCCGG
GCAGATTGGCGACTGGGCATGGGACACCGTCAGCCGCCTGTGTGGCACGGACGTGCTGACCGCGAC
CAACGCCTCAGGCGCGCCCACCTACCTGGCCTTCTATTACTTCCGCATCCGGGGCACGCCCGCGCTG
CATCCCGGCGCGCTGCGCTTCGGCGACACGCTGGACGTCACGTCGAAGGCGTACAACTTCGGCAGC
GAATCCGTCCTGACGGTGCACCGCATCTGCAAGACGGCGGAGGGCGGCGCTCCGGAGGCGGATGCC
TTCGGCCATGAAGAGCTGTACGAGCAGCCCCAGCCAGGCCGCATCTACGCGGAGACCTTCAACCGG
TGGATCACGCGCTCGGACGGCAAGTCGAACGAGAGCCTGATCAAGTCCTCGCCCGTGGGGTTCCAG
TACGCACACCTGCCGCTCTTGCCGGACGAATACTCGCCGCGGCGGGCCTATGGGGACGCGCGGGCG
CGGGGCACCTTTCACGATGTGGACTCCGCGGAGTACCGGCTGACCGTGGACCGCTTCCCGCTGCGCT
ACGCGGTGGACGTCATCCGGGACGTCAATGGGGTGGGGCTCATCTACTTCGCGTCGTATTTCTCGAT
GGTGGACTGGGCCATCTGGCAGCTGGCGAGGCACCAGGGACGCAGCGAGCAGGCCTTCCTGTCGCG
CGTGGTGCTGGACCAGCAACTGTGCTTCCTCGGCAACGCGGCGCTGGACACCACCTTCGACATCGAC
GTGCAGCACTGGGAGCGGGTGGGCGGCGGGGAAGAGCTGTTCAACGTGAAGATGCGCGAGGGCGC
GCAGGGCCGGGACATCGCCGTGGCGACGGTCAAGCTGCGCTTCGACGCCGCTTCGGAAGGAGGCCG
CCGTGGGTGAGCCGATGACAGACGAACAAATCCGCGGAGTCGTGCACCAGTCCATCGTGCGCGTCC
TGCCCCGCGTGCGCTCCAACGAGATTGCGGGCCACTTGAACCTCCGCGAGCTGGGCGCGGACTCCGT
GGACCGGGTCGAGATTCTCACGTCCATCCTGGACAGCCTGCGGCTGCAGAAGACGCCACTGGCGAA
GTTCGCCGACATCCGCAACATCGACGCGCTGGTGGCGTTCCTGGCCGGTGAGGTCGCGGGTGGCTG
AGCGGGTTCCCGGCGGAGTCGGCATCGAGGCCATCAACGCCTACGGCGGCGCCTCCATTCCGG
TGTTGGACTTGTTCCGGGGCCGGCGGCTGGACCCCGAAGCGATTCTCCAACCTGATGATGCAGGAGC
GCGGCGTCGCGCTGCCGTTCGAGGACCCCGTCACCAACGCGGTCAATGCGGCGCGGCCCATCCTGG
ACGCGATGTCGCCCGAGGCCCGGGAGCGCATCGAGCTCCTGGTCACCTCGAGCGAGTCCGGCGTGG
ACTTCAGCAAGTCCATCTCCTCGTATGCGCACGAGCACCTGGGGCTGAGCCGCCACTGCCGGTTCCT
GGAGGTGAAGCAGGCGTGTTACGCCGCCACCGGAGCGCTCCAGCTAGCGCTGGGCTACATCGCGTC
GGGCGTGTCACCGGGGGCCAAGGCCCTGGTGATTGCCACGGACGTGACGCTGGTGGACGAGAGCGG
TCTGTACTCCGAGCCGGCGATGGGCACCGGCGGCGTCGCCGTGCTGCTGGGCGACGAGCCGCGCGT
GATGAAGATGGACCTGGGAGCGTTCGGCAACTACAGCTACGACGTCTTCGACACCGCGCGGCCCTC
GCCGGAGATTGATATCGGCGACGTGGACCGGTCGCTCTTCACGTACCTGGACTGCCTCAAGCACAGC
TTCGCCGCGTATGGCCGCCGGGTGGACGGTGTCGACTTCGTGTCGACGTTCGACTACCTGGCGATGC
ACACGCCGTTCGCCGGACTGGTGAAGGCCGGGCACCGCAAGATGATGCGCGAGCTCACCCCGTGCG
ACGTGGACGAAATCGAAGCGGACTTCGGCCGGCGCGTGAAGCCGTCACTGCAGTACCCGAGTCTGG
TCGGGAACCTGTGCTCCGGCTCCGTGTACCTGAGCCTGTGCAGCATCATCGACACCATCAAGCCCGA
GCGGTCCGCTCGGGTGGGAATGTTCTCCTATGGGTCGGGTTGCTCGTCGGAGTTCTTCAGCGGCGTC
ATCGGCCCGGAGTCCGTGTCCGCGCTAGCTGGGTTGGACATCGGTGGCCACCTCCGGGGGCGCCGC
CAGCTCACGTTCGACCAATATGTCGAATTGCTGAAAGAGAACCTTCGCTGTCTGGTTCCAACGAAGA
ACCGGGACGTGGACGTGGAGCGCTACCTCCCGCTGGTGACGCGGACGGCGAGCCGCCCGCGCATGC

AGCGGCCGGGCCGTCAGCATCAGGCCAACGGGCAGACGACGGCCCACCTGGGGGCGGAAATCGCCT
CGCTGGCCGTGCCCGAAGGCGTCACCCCACAACTGTGGCGCTCGGCCACGTTTTCGGGGCAGGCCGC
ACTGGTGACCGTCCACGAGGCCTGGAACGCGGCGCGCCTCCAGGCCGTCCCCGGACACCGGATTGG
ATTGGTGGTGGGGGGCACCAACGTGCAGCAGCGCGACCTGGTGCTGATGCAAGACGCCTATCGCGA
GCGGGTGCCCTTTCTGCGGGCGGCCTACGGGTCGACCTTCATGGACACCGACCTCGTGGCCTCTGC
ACGCAGCAGTTCGCCATCCACGGGATGTCCTTCACGGTGGGAGGCGCATCGGCCAGTGGCCTGCTG
GCGGTCATCCAGGCCGCGGAGGCGGTGCTCTCAAGAAGGGTGGACGTTTGCATCGCCGTGGGGGCG
CTGATGGACGTCTCCTACTGGGAATGCCAGGGCCTGCGGGCCATGGGCGCGATGGGCACCGACCGG
TTCGCGCGGGAGCCGGAGCGTGCCTGCCGGCCCTTCGACCGGGAGAGTGATGGCTTCATCTTTGGAG
AGGCGTGCGGCGCCGTGGTGGTTGAGTCTGCGGAGCACGCTCGGCGACGCGGGGTGACTCCTCGCG
GCATCCTGTCGGGCTGGGCCATGCAGTTGGACGCGAGCCGCGGCCCGTTGTCGTCCATCGAAGGGG
AGTCGCAGGTGATTGGGGCTGCGCTGCGGCACGCGGACCTCGCGCCGGAGCGGGTGGACTACGTGA
ATCCTCACGGCAGCGGTTCGCGTCAGGGGGATGCCATCGAGCTGGGGGCCTTGAAGGCGTGCGGCC
TGACGCACGCCCGGGTCAACACCACGAAGTCCATCACCGGGCATGGCCTGTCCTCGGCGGGTGCCGT
GGGGCTCATCGCCACGCTGGTCCAGTTGGAGCAGGGCCGGCTGCACCCGTCCTTGAACCTGGTGGAC
CCGATTGATTCATCGTTCCGCTGGGTGGGGGCCACCGCGGAGGCCCAGTCCCTCCAGAACGCGCTGG
TGCTCGCCTACGGCTTCGGCGGCATCAACACCGCTGTCGCCGTGCGCCGGAGCGCCACGGAGAGCT
GACACGCCCATGCAAGCCGCTTCCCCTCCGCACCGCGACTACCAGACGCTCCGGGTCCGCTTCGAGG
CGCAGACCTGTTTTCTCCAGCTCCACCGGCCGGATGCGGACAACACCATCAGCCGCACGCTGATTGA
CGAGTGCCAGCAGGTGCTCACGTTATGTGAGGAGCACGCCACCACGGTGGTGCTCGAAGGCCTGCC
ACACGTGTTCTGCATGGGCGCGGATTTTCGAGCCATCCACGACCGGGTCGACGACGGCCGCCGGGA
GCAAGGCAACGCGGAGCAGCTGTACCGGCTGTGGCTGCAACTGGCGACAGGCCCCTACGTGACGGT
CGCCCATGTGCAGGGCAAGGCCAACGCGGGCGGCCTGGGCTTCGTCGCCGCGTGCGACATCGTGCT
GGCAAAGGCGGAGGTCCAGTTCAGTCTCTCCGAGCTGCTGTTCGGGCTGTTCCCCGCCTGCGTGATG
CCGTTCCTCGCCCGGCGAATCGGCATCCAGCGGGCGCACTACCTGACGCTGATGACGCGGCCCATCG
ACGCGGCCCAGGCGCTGAGCTGGGGGTTGGCGGACGCGGTGGACGCCGATAGCGAGAAGCTGTTGC
GGCTCCACTTGCGCAGGCTGCGGTGCCTGTCGAAGCCAGCGGTGACCCAGTACAAGAAGTACGCCT
CCGAGCTGGGCGGCCAGCTGCTCGCGGCCATGCCCCGGGCCATCTCCGCCAATGAGGCGATGTTCTC
CGACCGCGCCACGCTGGAAGCCATCCATCGCTACGTGGAGACAGGCCGACTCCCATGGAATCATG
ACGGAAGGAACGCCAATGGCGCCGGTGGTCACGCTCCATGAGGTGGAGGAGGGGGTGGCGCAGAT
CACCCTGGTGGATCGCGAGAACAAGAACATGTTCAGCGAGCAGCTCGTGCGCGAGCTCATCACCGT
GTTCGGCAAGGTGAATGGAAACGAGCGCTACCGCGCGGTGGTGCTCACCGGCTACGACACCTACTT
CGCGCTCGGCGGGACCAAGGCCGGCCTGCTGTCCATCTGCGACGGCATCGGCTCCTTCAACGTCACC
AACTTCTACAGCCTCGCGCTGGAGTGCGACATCCCGGTGATTTCCGCCATGCAGGGACATGGCGTAG
GCGGCGGGTTCGCGATGGGGCTGTTCGCGGACTTCGTGGTCCTGAGCCGGGAGAGCGTCTACACGA
CGAACTTCATGCGCTACGGCTTCACGCCGGGGATGGGCGCCACGTACATCGTGCCGAAGCGGCTGG
GGTACTCGCTCGGGCATGAGCTCCTGCTCAACGCCAGGAACTACCGCGGCGCCGACCTGGAGAAGC
GGGGCGTGCCTTTTCCGGTGTTGCCGCGCAAGGAAGTCTTGCCCCACGCCTACGAGATTGCGAGGGA
CCTGGCCGCGAAACCTCGGCTGTCGCTCGTGACGCTCAAGCGGCACCTGGTTCGCGACATCCGCCGA
GAGCTTCCGGACGTCATCGAGCGTGAGCTGGAGATGCACGGCATCACCTTCCATCACGACGACGTG
AGGAGGCGCATCGAGCAGCTGTTCCTCTGAGGCGCGCCCCTATGTTGAACCTGATCAACAACCACGC
ACACGGTTATGTGGTCACGCCCGTGGTCCTGGCCTGCAACGACGCTGGCCTGTTCGAACTCCTGCGG
CAGGGACCGAAGGACTTCGACCGGTTGGCGGAGGCATTGCGTGCCAACCGGGGACATCTGCGCGTC
GCGATGAGGATGTTCGAATCGCTCGGCTGGGTTCGCCGCGACGCGGATGACGTGTACGCGGTGACG
GCGGCGGCGGCCGCGCATCGGTCCTTCCCCCGCGAGGCGCAGTCGCTCTTCGCGCTGCCCATGGACC
GGTACCTGCGCGGGGAGGACGGCCTGTCCCTGGCGCCGTGGTTCGAGCGCTCTCGGGCGTCGTGGG
ATACCGATGACACGCTGGTGCGCGAGCTGCTCGACGGCGCCATCATCACGCCGCTGATGCTCGCGCT
GGAGCAGCGTGGGGGCCTCAAGGAGGCGAGGCGTCTGTCCGACCTGTGGTCCGGGGGGGATGGAA
GGGACACGTGCGTCCCCGAGGCCGTCCAACACGAGCTGGCCGGGTTCTTCTCCGCGCAGAAGTGGA
CGCGTGAGGACGCCGTCGACGCGGAGCTCACGCCCAAGGGCGCCTTCATCTTCGAGCGGGCATTGC
TCTTCGCCATCGTCGGCTCGTACCGGCCGATGCTGGCCAGCATGCCGCAGCTGCTCTTCGGTGACTG
CGACCAGGTCTTCGGGCGGGACGAAGCGGGCCACGAACTGCACCTGGACCGAACCCTCAACGTGAT
TGGGAGCGGCCACCAGCACCGGAAGTACTTCGCGGAGCTGGAGAAGCTCATCATCACCGTCTTCGA
TGCCGAGAACCTGTCGGCACAGCCGCGCTACATCGCGGACATGGGGTGCGGTGACGGCACGCTCCT
GAAGCGGGTGTATGAAACGGTGCTTCGGCACACGCGGCGGGGAAGGGCGCTCGACCGGTTTCCGCT
CACGCTCATCGCCGCGGACTTCAACGAGAAGGCGCTCGAAGCCGCTGGGCGGACGCTGGCCGGGTT
GGAGCACGTTGCCTTGCGCGCGGACGTGGCGCGGCCGGACCGTCTCATCGAGGACCTGCGGGCGCG
CGGGCTAGCCGAGCCTGAGAATACGCTGCACATCCGCTCGTTTCTCGACCACGACCGTCCCTACCAG
CCTCCCGCGGACAGGGCGGGGCTCCACGCCCGGATTCCGTTCGATTCGGTGTTCGTGGGCAAGGCG
GGCCAGGAGGTGGTTCCGGCGGAGGTGTTCCACAGCCTGGTGGAGCACCTCGAG